**Method and Apparatus for transforming a digital audio signal
and for inversely transforming a transformed digital audio
signal**

5   The invention relates to a method and to an apparatus for
transforming a digital audio signal from the time domain
into a different domain, and for inversely transforming a
transformed digital audio signal into the time domain.

10

Background

Known time domain to frequency domain or frequency domain to
time domain transformations used in codecs include the Dis-
15   crete Cosine Transform (DCT) or the Modified Discrete Cosine
Transform (MDCT). Both types of transformation have the dis-
advantage that they are costly in terms of required computa-
tional power since the computation involves multiplications
with a much higher precision than that of both the input and
20   the output values. E.g. in audio codecs, based on 16 bit in-
teger input samples and output values, in many cases the in-
ternal computations are carried out with at least 32 bit
fixed point or floating point precision. The input values
are multiplied with cosine values, which often are memorised
25   in look-up tables to reduce the processing power load. But
such tables consume valuable memory capacity which is pre-
cious in particular in embedded systems like audio players
or mobiles phones.
The Hadamard transformation does not use any such multipli-
30   cations but uses matrices consisting only of '+1' and '-1'
values. But using a Hadamard transform leads to reduced cod-
ing quality or increased bitrate.
The major advantage of the MDCT over the DCT is its lapped
nature, i.e. each input sample is transformed twice and each
35   output sample is the sum of two inverse transforms, which
has the effect that quantisation effects are averaged and

noise introduced by a is completely cancelled in the optimum
case. By subsampling following the overlapping the MDCT
transformed signal has as many samples as the input signal.
This feature is not feasible when using a Hadamard trans-
form. If an overlap of 50% is chosen, there are also 50%
more transformed samples, which fact contradicts the com-
pression goal and has strong drawbacks on transmission.

Invention

Most audio codecs transform their input data from time (or
space) domain to another domain (frequency domain), in which
compression and quantisation is carried out. However, DCT or
MDCT transformation is costly in terms of computation power
and memory.

A problem to be solved by the invention is to provide a
transform and a corresponding inverse transform which has
the advantages of MDCT but requires less computational
power. This problem is solved by the methods disclosed in
claims 1 and 2. Corresponding apparatuses that utilise these
methods are disclosed in claims 3 and 4, respectively.

The invention solves this problem by constructing a trans-
formation or inverse transformation which does not use any
multiplication apart from a single scaling, and which still
keeps the advantages of MDCT like overlapping and subsam-
pling. The related N*N full matrix is constructed by a com-
bination of two different N/2-rows N/4-columns sub-matrices
and reversed-column-order versions of these sub-matrices,
whereby the sub-matrices and thereby the (N/2)*N transforma-
tion matrix and the N*N full matrix contain '+1' and '-1'
values only.

The inventive transformation also represents a change be-

tween time domain and another domain. With proper overlap
and subsampling it is perfectly reconstructing. The inven-
tive transformation is very cheap in terms of computation
power and memory, since it does not use any multiplications
or high precision coefficient tables. Furthermore the inven-
tive transformation overlaps by 50% in the time domain which
reduces quantisation artefacts. At the same time it uses
subsampling by a factor of '2', i.e. a transformation of
length N samples results in N/2 transformed values. No sepa-
rate subsampling step is required. In combination with the
overlap a stream of L samples results in L transformed val-
ues (apart from a lead in and out) and in L inversely trans-
formed values. Advantageously, apart from the much lower
computing requirements the characteristics of the inventive
transform are very similar to that of the MDCT:
'smearing' of quantisation effects over the whole transfor-
mation length, 50% overlap so that quantisation artefacts
are averaged or even cancelled out, subsampling so that de-
spite a 50% overlap the number of transformed values is
equal to the number of input values.

In principle, the inventive methods are suited for:
transforming in an audio signal processor a digital audio
signal from the time domain into a different domain, includ-
ing the method steps:
- forming partitions of transform length N from said digi-
tal audio signal, which partitions overlap by N/2, wherein N
is an integer multiple of '4';
- performing a multiplication of a transform matrix Mh,
said transform matrix having a size of N/2 rows and N col-
umns, with each one of said partitions such that succeeding
transformed signal partitions are provided,
wherein said transform matrix is constructed in the form:
$$Mh = [a \quad lr(a) \quad b \quad lr(-1*b)] \; ,$$
wherein 'a' and 'b' are sub-matrices each having N/2 rows
and N/4 columns and including '+1' and '-1' values only,

and wherein said sub-matrices are linearly independent,
whereby said transform matrix multiplication outputs N/2
output values per N input values representing a subsampling
by a factor of '2', thereby forming a transformed digital

5 audio signal,
and for
inversely transforming in an audio signal processor a trans-
formed digital audio signal into the time domain, which
transformed digital audio signal was constructed by the

10 steps:
– forming partitions of transform length N from an original
digital audio signal, which partitions were overlapping by
N/2, wherein N is an integer multiple of '4';
– performing a multiplication of a transform matrix, said

15 transform matrix Mh having a size of N/2 rows and N columns,
with each one of said partitions such that succeeding trans-
formed signal partitions were provided,
wherein said transform matrix was constructed in the form Mh
= [a lr(a) b lr(-1*b)], wherein 'a' and 'b' were sub-

20 matrices each having N/2 rows and N/4 columns and including
'+1' and '-1' values only,
and wherein said sub-matrices are linearly independent,
whereby said transform matrix multiplication had output N/2
output values per N input values representing a subsampling

25 by a factor of '2', thereby having formed a transformed
digital audio signal,
including the method steps:
– performing a multiplication of an inverse transform ma-
trix invMh, said inverse transform matrix having a size of N

30 rows and N/2 columns, with each one of said transformed sig-
nal partitions such that succeeding inversely transformed
signal partitions of length N are provided,
wherein said inverse transform matrix invMH is constructed
by taking the left half of the inverse of a matrix

35 $$\begin{bmatrix} a & lr(a) & b & lr(-1*b) \\ b & lr(-1*b) & a & lr(a) \end{bmatrix},$$

wherein 'a' and 'b' are sub-matrices as defined above;
- assembling said inversely transformed signal partitions
in an overlapping manner so as to form an inversely trans-
formed digital audio signal, whereby said overlapping is of
5   size N/2,
and whereby the samples values of said inversely transformed
signal partitions, or the samples values of said inversely
transformed digital audio signal, or the values of said
transformed signal partitions are each scaled by multiplica-
10   tion with factor '1/N' or by a division by 'N' or by a cor-
responding binary shift operation.

In principle, the inventive apparatus for transforming a
digital audio signal from the time domain into a different
15   domain includes:
- means which form partitions of transform length N from
said digital audio signal, which partitions overlap by N/2,
wherein N is an integer multiple of '4';
- means which perform a multiplication of a transform ma-
20   trix Mh, said transform matrix having a size of N/2 rows and
N columns, with each one of said partitions such that suc-
ceeding transformed signal partitions are provided,
wherein said transform matrix is constructed in the form:
$$Mh = [a \quad lr(a) \quad b \quad lr(-1*b)] ,$$
25   wherein 'a' and 'b' are sub-matrices each having N/2 rows
and N/4 columns and including '+1' and '-1' values only,
and wherein said sub-matrices are linearly independent,
whereby said transform matrix multiplication means output
N/2 output values per N input values representing a subsam-
30   pling by a factor of '2', thereby forming a transformed
digital audio signal.

In principle, the inventive apparatus for inversely trans-
forming a transformed digital audio signal, which was con-
35   structed by the steps:
- forming partitions of transform length N from an original

digital audio signal, which partitions were overlapping by
N/2, wherein N is an integer multiple of '4';
- performing a multiplication of a transform matrix, said
transform matrix Mh having a size of N/2 rows and N columns,
5   with each one of said partitions such that succeeding trans-
formed signal partitions were provided,
wherein said transform matrix was constructed in the form Mh
= [a lr(a) b lr(-1*b)], wherein 'a' and 'b' were sub-
matrices each having N/2 rows and N/4 columns and including
10 '+1' and '-1' values only,
and wherein said sub-matrices are linearly independent,
whereby said transform matrix multiplication had output N/2
output values per N input values representing a subsampling
by a factor of '2', thereby having formed a transformed
15 digital audio signal,
into the time domain includes:
- means which perform a multiplication of an inverse trans-
form matrix invMh, said inverse transform matrix having a
size of N rows and N/2 columns, with each one of said trans-
20 formed signal partitions such that succeeding inversely
transformed signal partitions of length N are provided,
wherein said inverse transform matrix invMH is constructed
by taking the left half of the inverse of a matrix

$$\begin{bmatrix} a & lr(a) & b & lr(-1*b) \\ b & lr(-1*b) & a & lr(a) \end{bmatrix},$$

25 wherein 'a' and 'b' are sub-matrices as defined above;
- means which assemble said inversely transformed signal
partitions in an overlapping manner so as to form an in-
versely transformed digital audio signal, whereby said over-
30 lapping is of size N/2,
and whereby the samples values of said inversely transformed
signal partitions, or the samples values of said inversely
transformed digital audio signal, or the values of said
transformed signal partitions are each scaled by multiplica-
35 tion with factor '1/N' or by a division by 'N' or by a cor-
responding binary shift operation.

Advantageous additional embodiments of the invention are
disclosed in the respective dependent claims.

5   Drawing

Exemplary embodiments of the invention are described with
reference to the accompanying drawing, which shows in:
Fig. 1   Simplified block diagram for the inventive transfor-
10   mation in an audio signal processor, and for the in-
             ventive inverse transformation in an audio signal
             processor.

15   Exemplary embodiments

In Fig. 1a a digital input audio signal X is fed to a parti-
tioner PAR in which corresponding partitions x of length N
from signal X are formed. The partitions x are transformed
20   in a transform stage TRF, which gets transform matrix values
Mh from a memory MEM1, from the time domain into a different
domain, thereby providing the transformed output signal y.
Advantageously, the transformed x signal partitions are al-
ready subsampled by a factor of two so that no extra subsam-
25   pler is required. This signal can be encoded in a coder COD
including e.g. quantising, bit allocation and/or variable
length coding, whereby the resulting data rate is reduced
and encoding side information SI (for example encoding pa-
rameters) can be generated. The encoded audio signal is mul-
30   tiplexed in stage MUX with the side information SI, thereby
providing a signal T to be transferred.

In Fig. 1b the transferred signal T is fed to a demulti-
plexer stage DEMUX, which provides an encoded audio signal
35   together with side information SI to a decoder DEC. In DEC
the encoded audio signal is decoded using said side informa-

tion SI (for example encoding/decoding parameters), includ-
ing e.g. variable length decoding and/or inverse quantisa-
tion, and is thereafter fed as signal y' to an inverse
length-N transformer ITRF, which gets inverse transform ma-
trix values invMh from a memory MEM2, and which transforms
from said different domain back to the time domain. In stage
ASS the corresponding signal partitions x' of length N are
assembled in an overlapping manner thereby providing the
digital output audio signal X'.

The transformation of length N in transformer TRF is carried
out in an encoder such that in each case a corresponding
partition x of length N of a digital input audio signal X of
length L is transformed into a transformed signal y of
length N. This transformed signal y is transformed back in a
decoder in the inverse transformer ITRF to a corresponding
partition x' of an output signal X' such that X' equals X.
This is true, if the first N/2 and the last N/2 samples of
signal X are zero and L is an integer multiple of N/2. Since
each input signal X can be padded accordingly this means no
loss of generality.

The transformation length N must be an integer multiple of
'4', i.e. $n = N/4$, n and N being integer numbers.
The $(N/2)*N$ transformation matrix Mh has the form:

$$Mh = [a \; lr(a) \; b \; lr(-1*b)] \; ,$$

where 'a' and 'b' are sub-matrices having $2*n$ rows and n
columns consisting of only '+1' and '-1' values or elements.
E.g. "lr(a)" means that the columns or elements of sub-
matrix 'a' are reversed in order, i.e. $lr([1 \; 2 \; 3 \; 4])$ becomes
$[4 \; 3 \; 2 \; 1]$.

An $N*N$ full matrix MhFull is defined by:

$$MhFull = \begin{bmatrix} a \; lr(a) & b \; lr(-1*b) \\ b \; lr(-1*b) & a \; lr(a) \end{bmatrix}$$

The sub-matrices 'a' and 'b' are chosen such that their rows

are linearly independent from each other, i.e.

$$\text{rank}[\text{MhFull}] = N .$$

Advantageously, the inverse full matrix invMhFull is the in-
verse of full matrix MhFull scaled by N, so that the inverse
full matrix invMhFull consists of only '+1' and '-1' values,
too: $\text{invMhFull} = \text{inv}[\text{MhFull}] * N = N * [\text{MhFull}]^{-1}$.

The inverse transformation matrix invMh is formed by taking
the left half of inverse full matrix invMhFull. In 'Matlab'
software notation:

$$\text{invMh} = \text{invMhFull}[:, 1:(N/2)] .$$

Therein "[:," denotes that all rows are taken, ", 1:(N/2)]"
denotes that columns 1 to N/2 are taken.

An example transformation matrix for N=8 is:

$$\text{Mh} = \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 \end{bmatrix}$$

The corresponding full matrix is:

$$\text{MhFull} = \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix}$$

The corresponding inverse full matrix following multiplica-
tion with N is:

$$\text{invMhFull} = \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The corresponding inverse transformation matrix is:

$$
\text{invMh} = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 \\
1 & 1 & -1 & -1 \\
1 & 1 & 1 & 1 \\
-1 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 \\
-1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1
\end{bmatrix}
$$

With nTransforms = L/(N/2), i.e. the total length of input signal X divided by one half of the transform length equals the total number of transforms carried out on input signal X. For practical implementation, the value 'L' used does not correspond to the total length of audio signal X (e.g. the number of samples in 5 or 74 minutes) but to a usual audio coding frame length, e.g. in the range of 100 to 3000 samples. The transformation of the partitions x of input audio signal X from the time domain into a different domain is carried out as follows (in 'Matlab' software notation):

```
y = zeros(N/2, nTransforms);
for k = 0:(nTransforms-1)
    y(:, k+1) = Mh * x((1:N) + k*N/2);
end
```

The first line means that a matrix or a data field 'y' is generated which has N/2 rows and nTransforms columns, all of which are filled with '0' values.

According to the next line, k runs from '0' to (nTransforms-1) in the 'for' loop.

The third line expresses that the transformation matrix Mh is multiplied with an input signal vector x having the elements x(1+k*N/2) to x(N + k*N/2), each one of these multiplications yielding a vector having N/2 elements. The resulting (N/2)*nTransforms matrix is assigned to y.

The overlap of the transforms by N/2 is apparent. The transform coefficients of the overlapping partitions 'y' are subsampled by a factor of two.

The corresponding inverse transformation of the coefficients
of the partitions y of the transformed signal of the differ-
ent domain into corresponding partitions x' of the signal X'
in the time domain is carried out as follows (in 'Matlab'
software notation):

```
    x' = zeros(L, 1);
    for k = 0:(nTransforms-1)
        idx = (1:N) + k*N/2;
        x'(idx) = x'(idx) + invMh * y(:,k+1);
    end
    x' = x'/N
```

The first line means that a matrix or a data field x' is
generated which has L rows and a single column, all of which
are filled with '0' values.

According to the next line, k runs from '0' to (nTransforms-
1) in the 'for' loop.

The third line defines a parameter set idx having the ele-
ments (1+ k*N/2) to (N + k*N/2).

The fourth line expresses that the inverse transformation
matrix invMh is multiplied with a partial matrix of y con-
sisting of all rows of matrix y and column k+1 of matrix y,
whereby the resulting vectors each having N/2 elements are
summed up to form signal x'.

Since both the transform matrix Mh and the inverse matrix
invMh consist only of '+1' and '-1' values, the scaling in
the last line is the only multiplication (by factor '1/N'),
or division by 'N', in this transformation/inverse transfor-
mation, which multiplication or division can be implemented
as a shift operation in case N is a power of '2'. As an al-
ternative, the transformed input values of the inverse
transform can be scaled instead. Advantageously, all other
operations can be implemented as sums and differences.

By the overlapping, quantisation artefacts are averaged or
even cancelled. Following the inverse transform, the alias
introduced by the subsampling is also cancelled, i.e. a
'perfect reconstruction' is achieved.

As an alternative, the invention can be carried out with correspondingly transposed transform and inverse transform matices, i.e. matrix Mh has N rows and N/2 columns, whereas matrix invMh has N/2 rows and N columns.

The invention can be applied in audio coding/decoding, in audio data compression and in audio data transmission, storage and reproduction.